Attorney Docket No.   IBNR-014   PATENT

<u>IN THE UNITED STATES PATENT AND TRADEMARK OFFICE</u>

BOX PATENT APPLICATION
Assistant Commissioner for Patents
Washington, D.C. 20231

<u>NEW APPLICATION TRANSMITTAL</u>

Transmitted herewith for filing is the patent application of

Inventor(s):   John K. Renwick and Ross W. Callon

For (title):   APPARATUS AND METHOD FOR
                   FORWARDING DATA ON
                   MULTIPLE LABEL-SWITCHED
                   DATA PATHS

1.   **Type of Application** This new application is for
      a(n)

      ☒ Original (nonprovisional)

      ☐ Design

        ☐ Plant

      ☐ Divisional.

      ☐ Continuation.

      ☐ Continuation-in-part (C-I-P).

2.   **Papers Enclosed**

      __17__ Pages of specification

      __6__   Pages of claims

      __1__   Page of Abstract

      __3__   Sheets of drawings
              ☐ formal
              ☒ informal
      __1__   Page of Cover Sheet

☐   The **enclosed** drawing(s) are photograph(s), and there is also attached a "PETITION TO ACCEPT
     PHOTOGRAPH(S) AS DRAWING(S)." 37 C.F.R. 1.84(b).

3. **Additional papers enclosed**

☐ Preliminary Amendment

☐ Information Disclosure Statement (37 C.F.R. 1.98)

☐ Form PTO-1449 (PTO/SB/08A and 08B)

☐ Copies of cited references

☐ Declaration of Biological Deposit

☐ Submission of "Sequence Listing," computer readable copy and/or amendment pertaining thereto for biotechnology invention containing nucleoticle and/or amino acid sequence.

☐ Authorization of Attorney(s) to Accept and Follow Instructions from Representative

☐ Special Comments

☒ Other: Return Postcard.

4. **Declaration or oath**

☒ Enclosed

    ☐ Unexecuted

    ☒ Executed by

        ☒   inventors

        ☐   legal representative of inventor(s). 37 CFR 1.42 or 1.43.

        ☐   joint inventor or person showing a proprietary interest on behalf of inventor who refused to sign or cannot be reached.

        ☐   This is the petition required by 37 CFR 1.47 and the statement required by 37 CFR 1.47 is also attached. See item 12 below for fee.

☐ Not Enclosed

☐ Application is made by a person authorized under 37 C.F.R. 1.41 (c) on behalf of all the above named inventor(s).

        ☐   Showing that the filing is authorized.

5. **Assignment**

☒ An assignment of the invention to___IronBridge Networks, Inc._____.

    ☒ is attached. A separate ☐ "COVER SHEET FOR ASSIGNMENT (DOCUMENT) ACCOMPANYING NEW PATENT APPLICATION" or ☒ FORM PTO 1595 is also attached.

    ☐   will follow.

6. Certified Copy

Certified copy(ies) of application(s)

Country	Appln. no.	Filed
Country	Appln. no.	Filed
Country	Appln. no.	Filed

from which priority is claimed

☐ is (are) attached.

☐ will follow.

7. Fee Calculation (37 C.F.R. 1.16)

CLAIMS AS FILED

	Number filed		Number Extra	Rate	Basic Fee 37 C.F.R. 1.16(a) **$760.00**
Total Claims (37 CFR 1.16(c))	30	- 20 =	10	$ 18.00	$180.00
Independent Claims (37 CFR 1.16(b))	2	- 3 =	0	$ 78.00	
Multiple dependent claim(s), if any (37 CFR 1.16(d))			+	$260.00	

☐ Amendment cancelling extra claims is enclosed.

☐ Amendment deleting multiple-dependencies is enclosed.

☐ Fee for extra claims is not being paid at this time.

Filing Fee Calculation $ 940.00

8.   **Small Entity Statement(s)**

☒ Verified Statement(s) that this is a filing by a small entity under 37 CFR 1.9 and 1.27 is (are) attached.

☐ Status as a small entity was claimed in prior application _____, filed on_____ from which benefit is being claimed for this application under:

         35 U.S.C.   ☐ 119(e),

                        ☐ 120,

                        ☐ 121,

                        ☐ 365(c),

and which status as a small entity is still proper and desired.

☐ A copy of the verified statement in the prior application is included.

Filing Fee Calculation (50% of A, B or C above)

$ 470.00

9.   **Fee Payment Being Made at This Time**

☐   Not Enclosed

☐   No filing fee is to be paid at this time.
         (This and the surcharge required by 3 7 C. F.R. 1.16(e) can be paid subsequently.)

☒   Enclosed

     ☒   Basic filing fee $ 470.00

     ☒   Recording assignment
     ($40.00; 37 C.F.R. 1.21(h))
     (See attached "COVER SHEET FOR ASSIGNMENT
     ACCOMPANYING NEW APPLICATION".) $ 40.00

                 Total fees enclosed   **$ 510.00**

10. Method of Payment of Fees

☒ Checks in the amount of $ _470.00, $40.00_

☐ Charge Account No. 12-0448 in the amount of $ _____
A duplicate of this transmittal is attached.

11. Authorization to Charge Additional Fees

☒ The Commissioner is hereby authorized to charge the following additional fees during the entire pendency of this application to Account No. 12-0448.

☒ 37 C.F.R. 1.16(a), (f) or (g) (filing fees)

☒ 37 C.F.R. 1.16(b), (c) and (d) (presentation of extra claims)

☐ 37 C.F.R. 1. 1 6(e) (surcharge for filing the basic filing fee and/or declaration on a date later than the filing date of the application)

☐ 37 C.F.R. 1.17 (application processing fees)

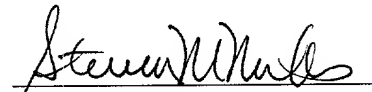☐ 37 C.F.R. 1.18 (issue fee at or before mailing of Notice of Allowance, pursuant to 37 C.F.R. 1.311(b))

12. Instructions as to Overpayment

☒ Credit Account No. 12-0448

☐ Refund

Respectfully submitted,

LAPPIN & KUSMER LLP

Steven M. Mills
Registration Number 36,610
Attorney for Applicants

Date _September 23, 1999_
Two Hundred State Street
Boston, MA 02109
617-330-1300
617-330-1311

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s):   John K. Renwick and Ross W. Callon

Filing Date: September 23, 1999

Title: APPARATUS AND METHOD FOR FORWARDING DATA ON MULTIPLE LABEL-SWITCHED DATA PATHS

CERTIFICATE OF MAILING UNDER 37 C.F.R.§ 1.10

"Express Mail" Mailing Label Number <u>EL024412536US</u> I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated below and is addressed to BOX PATENT APPLICATION, Assistant Commissioner for Patents, Washington, DC 2023

<u>9-23-99</u>
Date

Marcellus G. Green

BOX PATENT APPLICATION

Assistant Commissioner for Patents

Washington, DC 20231

Sir: <u>TRANSMITTAL LETTER</u>

Enclosed herewith for filing in the above-identified patent application please find the following listed items:

1.   New Application Transmittal;
2.   New Patent Application;
3.   Executed Declaration, Petition and Power of Attorney;
4.   Three (3) Sheets of Informal Drawings;
5.   Check in the amount of $470.00 to cover requisite fee;
6.   Executed Verified Statement;
7.   Assignment Recordation Form Cover Sheet - - PTO-1595;
8.   Executed Assignment;
9.   Check in the amount of $40.00 to cover assignment recordation fee; and
10.   Return Postcard.

In connection with the foregoing matter, please charge any additional fees which may be due, or credit any overpayment, to Deposit Account Number 12-0448. <u>A duplicate copy of this letter is provided for this purpose.</u>

Respectfully submitted,

LAPPIN & KUSMER LLP

Date: September 23, 1999

Two Hundred State Street

Boston, MA 02109

Telephone: (617) 330-1300

Facsimile: (617) 330-1311

Steven M. Mills

Registration Number 36,610

Attorney for Applicants

G:\IBNR\014\transapp.wpd

Applicant or Patentee: ___John K. Renwick and Ross W. Callon._____ Attorney's

Serial or Patent Number: __N/A_____ Docket No: ___IBNR-014_____

Filed or Issued: ___Filed herewith_____

For: __APPARATUS AND METHOD FOR FORWARDING DATA ON MULTIPLE LABEL-SWITCHED DATA PATHS_____

VERIFIED STATEMENT (DECLARATION) CLAIMING SMALL ENTITY STATUS (37 CFR 1.9 (f) and 1.27 (b)) -- SMALL BUSINESS CONCERN

I hereby declare that I am

☐   the owner of the small business concern identified below:

☒   an official of the small business concern empowered to act on behalf of the concern identified below:

NAME OF CONCERN   IronBridge Networks, Inc._____

ADDRESS OF CONCERN   55 Hayden Avenue, Lexington, Massachusetts 02173_____

I hereby declare that the above identified small business concern qualifies as a small business concern as defined in 13 CFR 121.3-18, and reproduced in 37 CFR 1.9(d), for purposes of paying reduced fees under section 41(a) and (b) of Title 35, United States Code, in that the number of employees of the concern, including those of its affiliates, does not exceed 500 persons. For purposes of this statement, (1) the number of employees of the business concern is the average over the previous fiscal year of the concern of the persons employed on a full-time, part-time or temporary basis during each of the pay periods of the fiscal year, and (2) concerns are affiliates of each other when either, directly or indirectly, one concern controls or has the power to control the other, or a third party or parties controls or has the power to control both.

I hereby declare that rights under contract or law have been conveyed to and remain with the small business concern identified above with regard to the invention, entitled:

___APPARATUS AND METHOD FOR FORWARDING DATA ON MULTIPLE LABEL-SWITCHED DATA PATHS_____

_____ by inventor(s)

John K. Renwick and Ross W. Callon_____ described in

☒   the specification filed herewith

☐   application serial no. _____ , filed _____

☐   patent no. _____ ,issued _____

If the rights held by the above identified small business concern are not exclusive, each individual, concern or organization having rights to the invention is listed below* and no rights to the invention are held by any person, other than the inventor, who could not qualify as a small business concern under 37 CFR 1.9 (d) or by any concern which would not qualify as a small business concern under 37 CFR 1.9 (d) or a nonprofit organization under 37 CFR 1.9 (e).

    *NOTE: Separate verified statements are required from each named person, concern or organization having rights to the invention averring to their status as small entities. (37 CFR 1.27)

FULL NAME ___NONE_____

ADDRESS _____

☐   INDIVIDUAL   ☐   SMALL BUSINESS CONCERN   ☐   NONPROFIT ORGANIZATION

FULL NAME _____

ADDRESS _____

☐   INDIVIDUAL   ☐   SMALL BUSINESS CONCERN   ☐   NONPROFIT ORGANIZATION

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28 (b))

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent to which this verified statement is directed.

NAME OF PERSON SIGNING   Kathleen Huber_____

TITLE OF PERSON OTHER THAN THE OWNER   Vice President_____

ADDRESS OF PERSON SIGNING   55 Hayden Avenue, Lexington, Massachusetts, 02173_____

SIGNATURE _____Kathie Huber_____ DATE __9/21/99__

G:\IBNR\014\verifistat.wpd

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

PATENT APPLICATION

of

John K. Renwick

and

Ross W. Callon

for

APPARATUS AND METHOD FOR FORWARDING DATA ON MULTIPLE

LABEL-SWITCHED DATA PATHS

APPARATUS AND METHOD FOR FORWARDING DATA ON MULTIPLE LABEL-SWITCHED DATA PATHS

Background of the Invention

In digital computer networks such as the Internet, collections of data, referred to as "datagrams," are typically transferred from node to node over the network in packets. Each packet of data typically includes a header portion and a data portion. In accordance with the common Internet protocol (IP), the header portion typically includes a 32-bit source identifying portion which identifies the source node that originated the packet and a 32-bit destination identifying portion which identifies the destination node to which the packet is ultimately to be transferred.

At each node, a router is used to forward the packet to the next node in the path toward the destination node. When a router receives a packet, it examines the destination address in the packet header. It then searches its locally stored routing table to determine the next node to which the packet should be transferred in order to ensure that it will reach its destination, typically along the shortest possible path. The router then forwards the packet to the next node identified in the routing table. This process continues at each successive node until the destination node is reached.

Another technique used to forward packets which can be implemented in IP routers is referred to as multiprotocol label switching (MPLS), or simply label switching. In MPLS, a path or route, referred to as a label-switched path (LSP), through a label-switched network or subnetwork is created before actual data traffic is forwarded over the network. The LSP directs the forwarding of datagrams from a given ingress point to a given egress point in the bounded

network or subnetwork.

In general, any pair of network nodes may be directly connected by more than one physical link. These physical links are equivalent to each other in the sense that a packet may be sent over any of them without affecting the way it is delivered to the ultimate destination. A set of such links comprises a single "logical link." This configuration is referred to herein as "multiple parallel links." In traditional IP forwarding, data traffic for a particular source and destination, i.e., a "flow," should always travel along the same link to ensure that data packets are not misaligned in time when they arrive at the destination node. Therefore, a router must select one of multiple links for transmission of data packets to the next link, and, for each source/destination pair, it must select the same link to ensure proper alignment of data at the destination. The router typically performs an analysis of the packet header contents to assign each packet to a physical link. Usually this involves a hash function of the 5-tuple of fields in the IP header (source IP address, destination IP address, protocol, source port number, destination port number) or a subset of these fields, such as source IP address and destination IP address. A hash function is designed to perform a computation on one or more data words and return a unique data word of shorter length. For example, a hash function performed on two 32-bit IP addresses may divide the combined 64-bit word by a constant and return as a result the value of the remainder in fewer bits, e.g., five. Other hash procedures include the use of a cyclic redundancy check (CRC) and the use of a checksum. Each time a hash procedure is performed on the same initial values, the same result is obtained. Therefore, deriving the link assignment from such a hash ensures that all packets with the same field contents take the same physical link, and thus all packets of a given application flow also take the same link. This ensures that the existence of

multiple links does not contribute to the risk of misordered packets within a flow.

In MPLS packet forwarding, MPLS can be used to route some traffic over a path other than the shortest one in order to relieve congestion on some paths. In a label switching network, paths are set up in advance for given sets of traffic that are to be forwarded along the same path. A set of traffic that is to follow the same paths is commonly referred to as a forwarding equivalence class (FEC). The ingress router can create label-switched paths (LSPs) to each of the other edge routers to which it expects to send traffic, and it can make all the decisions about the routes taken by the packets it forwards. It can do so by asking each downstream router in the path to assign a label value to the path. Using a signaling protocol such as RSVP or LDP, the ingress node sends a path setup request signal along the desired path to request labels from each of the nodes in the path. When the signal reaches the egress node, the egress node begins the process of allocating labels for the path. The egress node sends its allocated label back to the next preceding node, which stores the label and generates its own label for the traffic and transmits that label back to its next preceding node. This continues until all of the nodes along the path have assigned labels for the given FEC traffic.

When the ingress router desires to send a packet along a path, it places a small header, referred to as the MPLS header, on the front of the packet. The MPLS header contains the label value assigned by the next router in the path. The router then forwards the packet with the new MPLS header to the next router, which removes the label from the packet and replaces it with the label assigned by the next router. The packet is then forwarded to the next router. This continues until the packet has been forwarded over the entire LSP, i.e., it reaches the egress router. The egress router knows that it is at the end of the LSP, so it removes the

MPLS header from the packet and forwards the packet on the network using the traditional IP destination address lookup.

In MPLS, the labels assigned by the nodes identify the route to be taken from node to node along the LSP, but they do not readily provide the ability to distinguish multiple physical parallel links within the route or path. One reason for this is that it is difficult to compute an IP address hash at each node or "hop," as is done in IP forwarding, because the IP header is "hidden" behind the MPLS header. Hence, the hash procedure, which may be usable to select one of many paths, is not readily adaptable to MPLS. One of the features of MPLS is that the LSP carries opaque traffic, that is, the end points of the LSP know what protocol is carried, but the interior nodes do not. This feature could allow MPLS to be used to create virtual private networks that carry other protocols as well as IP. This means that an interior node cannot reliably compute an address hash, because it does not know anything about the contents of the packet behind the first MPLS header. Furthermore, even if it is known that all packets carry IP datagrams, there is a problem knowing where the IP header is located within the packet. This is because MPLS permits a stack of label headers to be added to the front of a packet. The label header format includes a single bit indicating the last label before the encapsulated protocol begins, meaning that the position of the encapsulated header must be found by examining the label headers, one-by-one, until the last one is found.

Conventional MPLS suffers a drawback in that, with conventional MPLS, it is relatively inefficient to set up multiple parallel paths to distribute traffic over multiple parallel physical links. In particular, each individual path needs to be independently signaled, i.e., set up. This signaling becomes particularly inefficient when there are a large number of potential parallel paths.

Summary of the Invention

The present invention solves the above problems and drawbacks of the prior art by providing a technique for allocating multiple paths in a route that is defined from node to node in a label switching network and a technique for distributing traffic of an FEC over multiple paths while ensuring that traffic of individual flows is not sent over different paths. The invention is directed to an apparatus and method for forwarding data from a source to a destination over a network which includes a subnetwork within the network, which in one embodiment is a label-switching subnetwork. The subnetwork includes a plurality of subnetwork nodes connected by a plurality of subnetwork links. The subnetwork nodes include an ingress node and an egress node coupled to the source and destination, respectively. At least one pair of subnetwork nodes is connected by a plurality of subnetwork links, i.e., parallel physical links. The subnetwork nodes and links define a plurality of subnetwork paths between the ingress node and the egress node. A response request signal is forwarded from the ingress node to the egress node along a route through a subset of subnetwork nodes between the ingress node and the egress node. The response request signal requests a response from each node along the route. The response signals sent by the nodes define or allocate a plurality of paths within the route between the ingress node and the egress node.

In general, the subnetwork is a label-switching network within a larger network such as the Internet which forwards data using the Internet protocol (IP). The ingress node can be coupled to and receive packets from multiple source nodes and/or nodes interposed between the source nodes and the ingress node. Likewise, the egress node can be coupled to and transmit packets to multiple destination nodes and/or nodes interposed between the egress node and the

destination nodes. In general, packets can be forwarded to the ingress node of the subnetwork using IP packet forwarding. The ingress node router can attach an MPLS header to the packet and forward it along the subnetwork using label switching in accordance with the invention to the egress node router. The egress node router then removes the MPLS header from the packet and forwards the packet to the next node on the way to the destination node using IP forwarding.

The data for each source/destination pair will be assigned a single path within the route. As a result, misalignment of data at the egress node is avoided.

In one embodiment, the response signals sent by the nodes simultaneously define the multiple paths in accordance with the invention. The response signal can include a label word which defines a number of data bits for the label. A certain predefined portion of the bits of the label word define the route to be used such that packets are forwarded along the correct path from node to node. The remaining bits are used to define multiple paths within the route allocated by the nodes to carry data. For example, in one embodiment, a complete label word is twenty bits long. To establish 32 paths, for example, the first fifteen bits of the response label word can be used by a node to identify the route. The remaining five bits can be used to select one of 32 possible paths within the route. Hence, in the response label word transmitted by a node, the first subset or group of bits, e.g., fifteen bits, identifies the route. The remaining, e.g., five, bits are "don't cares."

The ingress router can select one of the available allocated paths based on the source and destination of an arriving packet. For example, the ingress router can perform a hash operation on the IP source and destination fields in the IP header of the packet to produce a unique word of the same number of bits as the number of "don't care" bits in the response label. The resulting word can then be

used to select one of the paths. In the example above, where five bits are used to select one of 32 paths, a hash operation can be performed on the 32-bit source and destination IP addresses to produce a unique five-bit word, which is then used to select one of the 32 possible allocated paths.

In actuality, each node router need not actually transmit the don't care bits back up the route to the ingress router. Only the bits used to identify the route need be transferred. The ingress router can perform the necessary operations for selecting one of the allocated paths. Since the system provides for selection of a single path for a given source/destination pair via the hashing procedure, it is ensured that misalignment of data at the egress node is avoided.

The approach of the invention provides numerous advantages. For example, label switching, which is much faster and more efficient than IP forwarding, can be used efficiently in an environment with multiple parallel links. The system allows multiple paths to be set up simultaneously instead of requiring that each path be set up individually. This saves considerable processing time, which leads to improved network operation, particularly with respect to reduced time to set up or adjust paths. This also allows a bundled set of paths to be handled with reduced control resources, when compared to the resources required to set up paths individually.

Brief Description of the Drawings

The foregoing and other objects, features, and advantages of the invention will be apparent from the following more particular description of preferred embodiments of the invention, as illustrated in the accompanying drawings in which like reference characters refer to the same parts throughout the different views. The drawings are not necessarily to scale, emphasis instead being placed

upon illustrating the principles of the invention.

FIG. 1 is a schematic block diagram of a network 10 which can implement a packet forwarding technique in accordance with one embodiment of the invention.

5

FIG. 2 contains a detailed block diagram of the subnetwork shown in FIG. 1.

FIG. 3 is a schematic block diagram which illustrates a variation on the subnetwork of FIG 2 in which pairs of subnetwork nodes can be connected by multiple links.

10

15

Detailed Description of Preferred Embodiments

FIG. 1 is a schematic block diagram of a network 10 which can implement a packet forwarding technique in accordance with one embodiment of the invention. Referring to FIG. 1, the network 10 can include multiple nodes 16 connected by links 24 which carry data packets from node to node on the network 10. As shown, the network includes multiple source nodes 12 and destination nodes 14 connected to the network nodes 16. A source node 12 can forward data packets over the network 10 to a destination node 14. It will be understood that the source and destination nomenclature is used as a convention to describe the

20

direction of data flow. In conventional networks, all nodes can typically serve as a source node or a destination node, depending upon whether it is sending or receiving data.

In accordance with the invention, the network 10 includes a subnetwork 22 over which packets can be transferred en route from a source node 12 to a

25

destination node 14. Packets enter the subnetwork 22 through a node X(26) and exit the subnetwork 22 through a node Y(28). In accordance with the invention,

the network 10 is primarily an IP network which transfers packets between the nodes 16 using the common IP packet forwarding protocol. Each node includes a router with a stored routing table which allows the router to determine the node to which a packet should be forwarded based on a table lookup of the destination

5   address stored in the IP header of the packet. In accordance with the invention, the subnetwork 22 does not forward data packets using the IP protocol. It is an MPLS label switching network which forwards packets using label switching. Node X forwards packets into the subnetwork 22 using IP forwarding protocol. The packets are forwarded through the subnetwork 22 as described below in detail

10   using a label switching technique in accordance with the invention and are transferred out of the subnetwork 22 to node Y using IP packet forwarding. From there, the packets continue on to their designated destination nodes 14 using IP forwarding.

    FIG. 2 contains a detailed block diagram of the subnetwork 22 shown in

15   FIG. 1. The subnetwork 22 includes multiple nodes 17 connected by links 19. A datagram or packet from node X(26) to node Y(28) enters the subnetwork 22 at ingress node router A and exits through egress node router I. In traditional IP forwarding, each router in the network would independently compute the shortest paths for packets to all known destinations, and each router would forward each

20   packet by looking up its destination address in a forwarding or routing table. For a packet from X to Y, the shortest path would be A-B-G-I.

    Thus, traditional node-by-node or hop-by-hop IP routing would try to make every packet take the shortest path from entry to exit in the subnetwork. This is efficient, but it might leave some network resources relatively

25   underutilized, while causing some other resources to become overloaded. In FIG. 2, the path A-B-C-D-I is an alternative path for packets from X to Y that, while

longer than the path selected by hop-by-hop routing, could still be used to carry some of the load. This might help to relieve congestion on the path B-G-I.

MPLS can be used to route some of the traffic over a path other than the shortest one. In order to increase the degree of control of network path utilization, MPLS paths can be managed from a single point rather than hop-by-hop. For instance, the ingress router A can create label-switched paths (LSPs) to each of the other edge routers to which it expects to send traffic, and it can make all the decisions about the routes taken by the packets it forwards.

An ingress router A creates a label-switched path by asking each downstream router in the path to assign a label value to the path. In one implementation, a label is an integer value small enough to be used as an index value in a table lookup. As an example, it is assumed that ingress router A wishes to send set up a label-switched path A-B-C-D-I. Using a signaling protocol, such as RSVP or LDP, it sends a path setup request along that path. The routers in the path assign, for example, the following label values to it: B assigns 4, C assigns 20, D assigns 38, and I assigns 9. It should be noted that these label values are selected for purposes of illustrating an example of the invention. In practice, the values selected by the nodes in a path may in general be different each time that path is created. When the response comes back to A, the path is set up. Now if A wishes to forward a packet along that path, it places an MPLS header on the front of the packet containing the value assigned by B (4) and sends the packet on a link to B. B, receiving a packet labeled 4, knows it must froward the packet on a link to C with the new label value 20. It replaces the 4 with 20 in the label header and sends the packet to C. Each router in the path replaces the label value with the new one assigned by the next-hop router until the packet arrives at I, bearing the label value 9. Router I knows that it is the end of the path for packets labeled 9,

so it strips the label header from the packet and forwards it toward Y using the traditional IP destination address lookup.

In MPLS packet forwarding, each router stores a label map which defines how data with a particular label is to be routed through the network. A label map includes a table which includes an entry for the input circuit identifier, i.e., the identifying label for the router receiving the packet for forwarding, the output link, i.e., the link over which the packet should be forwarded out of the present node, and an output circuit identifier, i.e., the label identifier for the next succeeding node in the path to which the packet is to be forwarded. When a router at a node receives a packet, it examines the label map to determine the link on which the packet is to be forwarded to the next node. It discards the label that was on the packet when it arrived and replaces it with the label for the next node, such that the next node can use its own label map to forward the packet to the next succeeding node.

FIG. 3 is a schematic block diagram which illustrates a variation on the subnetwork 22 of FIG 2. In the subnetwork 122 of FIG. 3, multiple links 19 can be used between nodes 17 of the subnetwork 122. The multiple links are managed in such a way that the existence of multiple equivalent links between any pair of routers is known only to the link layer. In traditional IP forwarding, they appear as if they were a single link. Under IP forwarding, in the subnetwork 122 of FIG. 3, router A would look up the destination address in a packet for a destination coupled to node Y. The router A decides that the shortest path to Y follows one of the links from A to B. It must select one of five equivalent links for transmission of each packet in such a way as to balance the traffic load over all the available links. There are many ways of making this selection, but it is important to ensure that all packets from X to Y that belong to the same flow take

the same link. If they took different links, there is a danger that they would not arrive at Y in the same order that X sent them. Such misordering of packets can cause TCP protocol implementations to react as if packets were lost, resulting in unnecessary retransmission and reduced throughput. Thus router A performs an analysis of the packet header contents to assign each packet to a physical link. Usually this involves a hash function of the 5-tuple of fields in the IP header (source IP address, destination IP address, protocol, source port number, destination port number) or a subset of these fields, such as source IP address and destination IP address. Deriving the link assignment from such a hash ensures that all packets with the same field contents take the same physical link, and thus all packets of a given application flow also take the same kink. This ensures that the existence of multiple links does not contribute to the risk of misordered packets within a flow.

Using traditional MPLS to forward packets through the subnetwork of FIG. 3 presents certain problems. For example, using MPLS to route traffic over the path A-B-G-I in FIG. 3 is considered. The desire is to distribute the traffic evenly across all the links, without producing out-of-order delivery. In this case, it is difficult to compute an IP address hash at each node because the IP header is hidden behind the MPLS header. In accordance with the invention, optimal distribution of the traffic over the multiple links in this path using MPLS is achieved by creating multiple label-switched paths (LSPs) between nodes. In one embodiment, the number of LSPs needed to create the path from the ingress router A to the egress router I is equal to the least common multiple (LCM) of the number of links on each individual hop. For example, for the path A-B-G-I, the number of links on the A-B hop is 5, the number of links on the B-G hop is 3, and the number of links on the G-I hop is 2. Therefore, the number of LSPs generated

for the A-B-G-I path should be a minimum of the least common multiple (LCM) of 5, 3 and 2; LCM (5,3,2) = 30. The A-B hop would divide the 30 LSPs into five groups of six; the B-G hop would divide the 30 LSPs into three groups of ten; and the G-I hop would divide the LSPs into two groups of fifteen.

Whereas LCM computes the minimal number of paths needed to balance the distribution of paths over a plurality of physical links, this minimal number may be much fewer that the most desirable number. For example, if the number of hops from A to B and B to G in FIG. 3 where changed to 8, LCM (8, 8, 2) gives 8 as the number of paths needed to ensure equal distribution of paths over links on each hop. But if one of the links from B to G were to fail, then LCM (8, 7, 2) = 56. In this situation it becomes impossible to balance the load over the remaining links using only eight switched paths. In this example, a number of paths on the order of eight times the LCM may be required to ensure the load can still be a balanced after one or more link failures.

In accordance with the invention, these multiple LSPs are generated simultaneously in response to a single request signal from the ingress router A sent along the path. In accordance with one embodiment of the invention, a mask value is added to the label value when setting up a LSP. By way of example, the MPLS standard uses a 20-bit label. The mask value added according to one embodiment is also a 20-bit number in which a bit is set (1) to indicate that the corresponding label value bit is important in determining the route that packets will take, and clear (0) to indicate that it is not, i.e., that it is a "don't care" bit. The effect is to simultaneously set up a number of LSPs that take the same router-to-router path, but that do not necessarily use the same links between the routers when multipath links are present. The number of LSPs simultaneously created and maintained can be 1, 2, 4, 8, 16, 32, 64, etc., depending on the number of bits

that are zero in the mask value. In general, if the number of zero bits is n and the number of paths set up is N, then, in one embodiment, $N=2^n$. These multiple LSPs are referred to herein as an "LSP bundle."

A number of alternative implementations for the mask value of the invention are possible. In one embodiment, the mask value can be transmitted as part of the path setup message. Alternatively, the mask value can be configured in each router in the path. The mask value can be represented as a 20-bit field with ones in the mask position. Alternatively, the mask value can be represented by an integer indicating the number of leading one bits (or trailing zero bits) in the mask. Use of an integer value implies that the mask is a contiguous string of one bits starting in the left-most bit position. It will be understood that other configurations are possible. Also, the bit value interpretations, i.e., one or zero, can be reversed.

The operation of setting up the LSP bundle is analogous to setting up a single path as described above. Assuming by way of example only that a 64-path LSP bundle is to be created for the route A-B-G-I in FIG. 3, in one embodiment, router A sends a path setup request along that path, sending the binary mask value 1111 1111 1111 1100 0000. Each router in the path assigns 64 label values for the 64 bundled LSPs that are being created. The label values are assigned so that they all share the same bit pattern in the bits corresponding to "1" bits in the mask value. One such set of assignments is shown by way of example only in Table 1 below.

Router	Label Range Assigned
B	$0000\ 0000\ 0001\ 0000\ 0000_2$ - $0000\ 0000\ 0001\ 0011\ 1111_2$ (256_{10}-319_{10})
G	$0000\ 0000\ 0101\ 0000\ 0000_2$ - $0000\ 0000\ 0101\ 0011\ 1111_2$ (1280_{10}-1343_{10})
I	$0000\ 0000\ 0010\ 0100\ 0000_2$ - $0000\ 0000\ 0010\ 0111\ 1111_2$ (576_{10}-639_{10})

Table 1.

Each router assigns the 64 labels assigned by the downstream router to the links to the next hop so that each link carries approximately the same number of paths. For example, router A could divide router B's 64 labels into four groups of 12 and one group of 16, assigning labels 256 through 267 to one link, 268 through 279 to the next link, 280 through 291 to the next link, 292 through 303 to the next link, and 304 through 319 to the fifth link.

Router A is the entry or ingress to the LSP. It is the last point at which the IP header contents are visible. In one embodiment, router A assigns incoming traffic to one of the LSPs using the IP header. In one embodiment, it computes an IP address hash for each packet and uses the hash value to assign each packet to one of the 64 LSPs. In this particular example, the hash operation returns a 6-bit word used to select one of the 64 LSPs. The hash operation used to compute the hash value can be of the type described in, for example, The Art of Computer Programming, Volume 3, "Sorting and Searching," by Donald E. Knuth, published in 1973 by Addison-Wesley, pages 512-513, or other suitable hash operation. The hash operation can include performing a division on the IP source and destination addresses, such as by dividing them by a constant and keeping the remainder as the hash result. Alternatively, the hash can include a cyclic redundancy check (CRC) or a checksum. It will be recognized that other hash procedures can be used in accordance with the invention. The result, given a sufficiently large aggregation of application flows along the route, is a statistically even distribution of traffic across the available links to the egress router I. It should be noted that the hash operation need not be performed on the address fields in the IP header of the packets. Alternatively, the hash can be performed on the protocol field or some other information which identifies the packet as being part of a particular flow between a source node and a destination node.

In accordance with the invention, multiple LSPs can be merged, that is, packets can arrive at a router on different links or with different labels and go out on the same link with the same label. There are three cases considered with respect to merging of

LSPs. The first case is the one in which an LSP bundle is joined to a single LSP. In this case, there are two possibilities: either the single LSP can become a bundle, or all the LSPs in the bundle can merge into the single LSP. This choice is a matter of network administration, and it does not increase the possibility that packets will be misordered. The second merging case is the case in which one LSP bundle joins another LSP bundle. In this case, there are three choices for merging: the number of LSPs proceeding forward from the merge point can be the number of LSPs in the original bundle, or it can be the number in the bundle that joined it, or it can be the sum of the two. When mapping one LSP into another, it is only necessary to ensure that packets entering on one LSP do not somehow exit on more than one, since that would cause packet misordering. The third merging case is the case in which a single LSP joins a bundle. In this case, the single LSP must join just one of the LSPs in the bundle.

It is desirable that all routers in the path of the bundled LSP support the mask in the signaling protocol. If they do not, it is still possible to set up bundled LSPs in the routers that support them using a static configuration of the mask length. The mask length should be configured to 20, for example, i.e., exactly one LSP in a bundle, on links toward routers that do not support bundling. If a single LSP is configured to branch into a bundle at some point in the path, it is necessary to look inside the packet, determine if and where an IP header exists within, and hash its contents to assign each packet to one of the bundled LSPs.

It is also possible to signal bundled LSPs along such a heterogeneous path. There are at least two cases considered. The first case is that of a path with a set of links from a router that uses bundled LSPs to one that does not. To illustrate this case, referring to FIG. 3, it is assumed that router A supports bundling and router B does not. When router A receives a bundled LSP setup request for a path leading to B, it can replicate this as it propagates it to B. Each LSP in the bundle generates an individual request to B, so that the number of LSPs in the bundle is maintained.

The second case is that of a path with a set of links from a router that does not use bundled LSPs to one that does. To illustrate this case, referring to FIG. 3, it is assumed that router A supports bundling and router B does not. However, in this case, it is also assumed that router B is setting up a group of paths that lead to router A.

5 When it receives the first request, A can pass it along to the next hop as a bundled request. As it receives subsequent requests from B for LSPs taking the same route, it can map them into the established bundle and send a successful reply to B, without propagating the request further downstream.

An MPLS path or "tunnel" can be configured to be carried inside a second

10 MPLS tunnel by adding a second MPLS header to the front of the packet. Both inner and outer tunnels can be bundled LSPs in accordance with the invention. Individual LSPs of the inner bundle can be mapped to those of the outer bundle. The inner label value implicitly carries the IP address hash information that was calculated at the original ingress. In effect, this same hash can determine LSP assignment into any

15 number of tunnels in a hierarchy.

While this invention has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the spirit and scope of the invention as defined by the appended claims.

20 What is claimed is:

CLAIMS

1   1.   A method of forwarding data over a network from a source node to a

2   destination node, comprising:

3   providing a subnetwork within the network having a plurality of

4   subnetwork nodes connected by a plurality of subnetwork links, the

5   subnetwork nodes including an ingress node and an egress node coupled to the

6   source node and the destination node, respectively, at least one pair of

7   subnetwork nodes being connected by a plurality of subnetwork links, the

8   plurality of subnetwork nodes and the plurality of subnetwork links defining a

9   plurality of subnetwork paths between the ingress node and the egress node;

10   forwarding a signal from the ingress node to the egress node along a

11   route through a subset of subnetwork nodes between the ingress node and the

12   egress node, said signal requesting a response from each node along the route;

13   and

14   receiving response signals from the nodes along the route, the response

15   signals defining a plurality of paths within the route between the ingress node

16   and the egress node.

1   2.   The method of claim 1 wherein the subnetwork comprises a label-switching

2   network.

1   3.   The method of claim 2 wherein the network comprises nodes which forward

2   data using Internet protocol node addresses.

1
2

4.   The method of claim 2 wherein each subnetwork node along the route allocates a plurality of labels for the plurality of paths along the route.

1
2
3
4
5

5.   The method of claim 2 wherein:

       the ingress node is coupled to a plurality of source nodes; and

       each source node coupled to the ingress node is associated with one of the plurality of paths along the route between the ingress node and the egress node.

1
2
3
4
5

1
2
3

6.   The method of claim 2 wherein:

       the egress node is coupled to a plurality of destination nodes; and

       each destination node coupled to the egress node is associated with one of the plurality of paths along the route between the ingress node and the egress node.

7.   The method of claim 2 further comprising associating each packet of data to be transferred from a particular source node to a particular destination node with one of the plurality of paths between the ingress node and the egress node.

1
2

8.   The method of claim 7 wherein the associating comprises performing a logical operation on information carried in each packet of data.

1
2

9.   The method of claim 8 wherein the logical operation comprises a hash operation.

1 10. The method of claim 8 wherein the logical operation is performed on an

2   address field in the packet of data.

1 11. The method of claim 8 wherein the logical operation is performed on a

2   protocol field in the packet of data.

1 12. The method of claim 2 wherein a response signal includes a label word which

2   defines a plurality of data bits, a first subset of the defined data bits being

3   associated with the route between the ingress node and the egress node and a

4   second subset of the defined data bits being associated with the plurality of

5   paths within the route.

1 13. The method of claim 12 wherein the data bits of the second subset of the

2   defined data bits are not assigned values by the node that generated the

3   response signal.

1 14. The method of claim 12 wherein the number n of data bits in the second subset

2   of the defined data bits determines the number N of defined paths within the

3   route.

1 15. The method of claim 14 wherein $N=2^n$.

1 16. The method of claim 1 wherein:

2   the ingress node is coupled to a plurality of source nodes; and

3   each source node coupled to the ingress node is associated with one of

4   the plurality of paths along the route between the ingress node and the egress

5   node.

1   17.   The method of claim 1 wherein:

2   the egress node is coupled to a plurality of destination nodes; and

3   each destination node coupled to the egress node is associated with one

4   of the plurality of paths along the route between the ingress node and the

5   egress node.

1   18.   The method of claim 1 wherein a subnetwork link between a pair of

2   subnetwork nodes is assigned to carry a plurality of the defined paths between

3   the ingress node and the egress node.

    19.   The method of claim 1 wherein the plurality of subnetwork links connecting

      the at least one pair of subnetwork nodes form a single logical link used in

      forwarding the data from the ingress node to the egress node.

    20.   An apparatus for forwarding data over a network from a source node to a

      destination node, comprising:

                     a subnetwork within the network having a plurality of subnetwork

      nodes connected by a plurality of subnetwork links, the subnetwork nodes

5   including an ingress node and an egress node coupled to the source node and

6   the destination node, respectively, at least one pair of subnetwork nodes being

7   connected by a plurality of subnetwork links, the plurality of subnetwork nodes

8   and the plurality of subnetwork links defining a plurality of subnetwork paths

9   between the ingress node and the egress node; and

10   a communication subsystem within the subnetwork for (i) forwarding a

11   signal from the ingress node to the egress node along a route through a subset

12   of subnetwork nodes between the ingress node and the egress node, said signal

13   requesting a response from each node along the route, and (ii) forwarding

14		response signals from the subnetwork nodes along the route, the response
15		signals defining a plurality of paths within the route between the ingress node
16		and the egress node.

1	21.	The apparatus of claim 20 wherein the subnetwork comprises a label-switching
2		network.

1	22.	The apparatus of claim 21 wherein the network comprises nodes which
2		forward data using Internet protocol node addresses.

1	23.	The apparatus of claim 21 wherein each subnetwork node along the route
2		allocates a plurality of labels for the plurality of paths along the route.

1	24.	The apparatus of claim 21 wherein:
2		the ingress node is coupled to a plurality of source nodes; and
3		each source node coupled to the ingress node is associated with one of
4		the plurality of paths along the route between the ingress node and the egress
5		node.

1	25.	The apparatus of claim 21 wherein:
2		the egress node is coupled to a plurality of destination nodes; and
3		each destination node coupled to the egress node is associated with one
4		of the plurality of paths along the route between the ingress node and the
5		egress node.

1	26.	The apparatus of claim 21 wherein a response signal includes a label word
2		which defines a plurality of data bits, a first subset of the defined data bits
3		being associated with the route between the ingress node and the egress node

1　　　and a second subset of the defined data bits being associated with the plurality

2　　　of paths within the route.

1　27.　The apparatus of claim 26 wherein the data bits of the second subset of the

2　　　defined data bits are not assigned values by the node that generated the

3　　　response signal.

1　28.　The apparatus of claim 26 wherein the number n of data bits in the second

2　　　subset of the defined data bits determines the number N of defined paths within

3　　　the route.

　29.　The apparatus of claim 20 wherein a subnetwork link between a pair of

　　　subnetwork nodes is assigned to carry a plurality of the defined paths between

　　　the ingress node and the egress node.

　30.　The apparatus of claim 20 wherein the plurality of subnetwork links

　　　connecting the at least one pair of subnetwork nodes form a single logical link

　　　used in forwarding the data from ingress node to the egress node.

APPARATUS AND METHOD FOR FORWARDING DATA ON MULTIPLE LABEL-SWITCHED DATA PATHS

Abstract of the Disclosure

An apparatus and method for forwarding data on a network are described. A label-switching subnetwork within the network includes an ingress node and an egress node coupled to source and destination nodes, respectively, on the network. The ingress node sends a signal along a route within the subnetwork through a plurality of subnetwork nodes to the egress node. In response, the subnetwork nodes transmit response signals back along the route toward the ingress node which define the route through the subnetwork and simultaneously allocate a plurality of paths within the route. A single path can be selected for forwarding of data packets associated with a source/destination pair, ensuring that data packets arriving at the destination are not misaligned.

5

10

G:\IBNR\014\PATAPP2.WPD
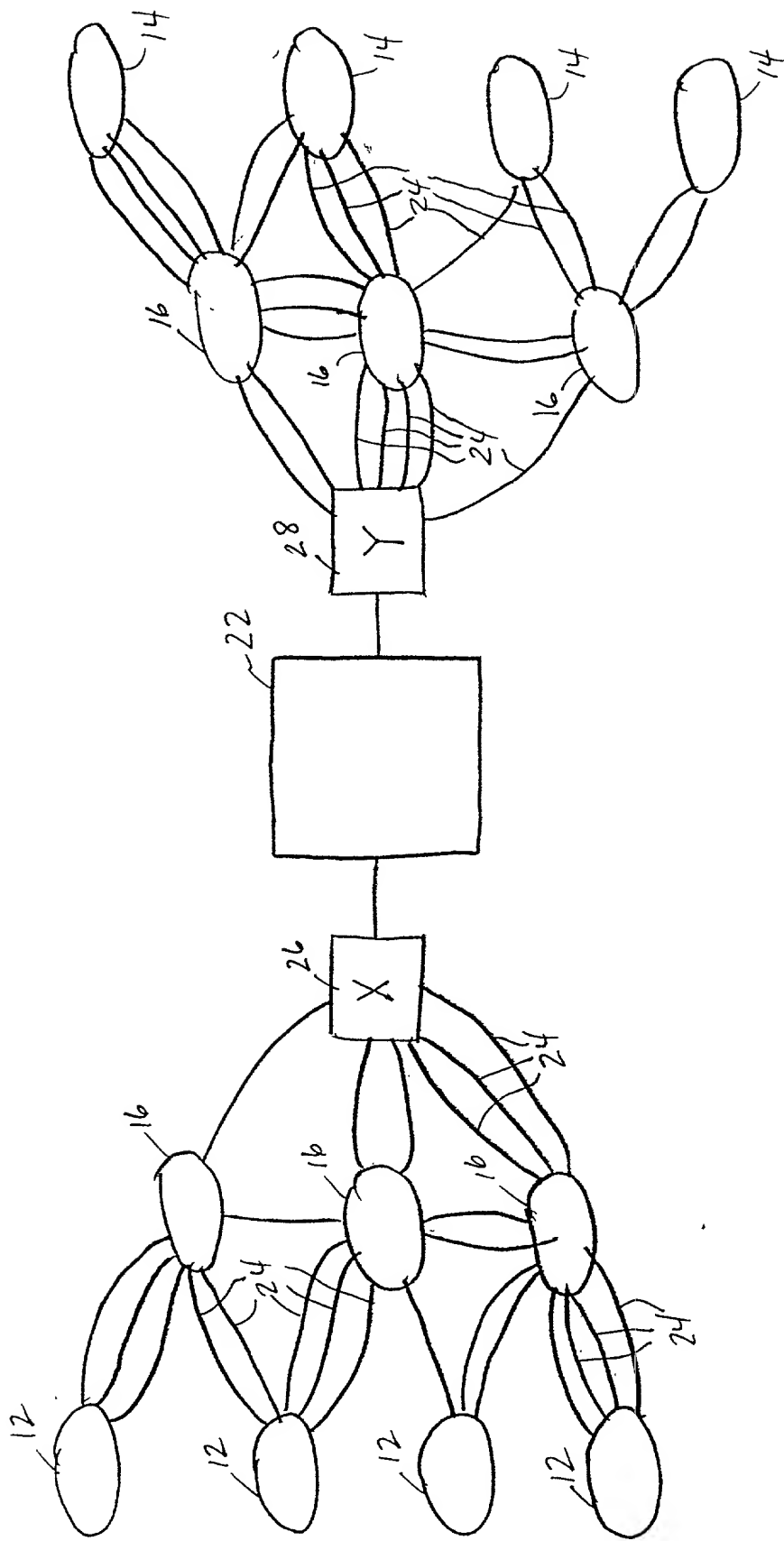
FIG. 1
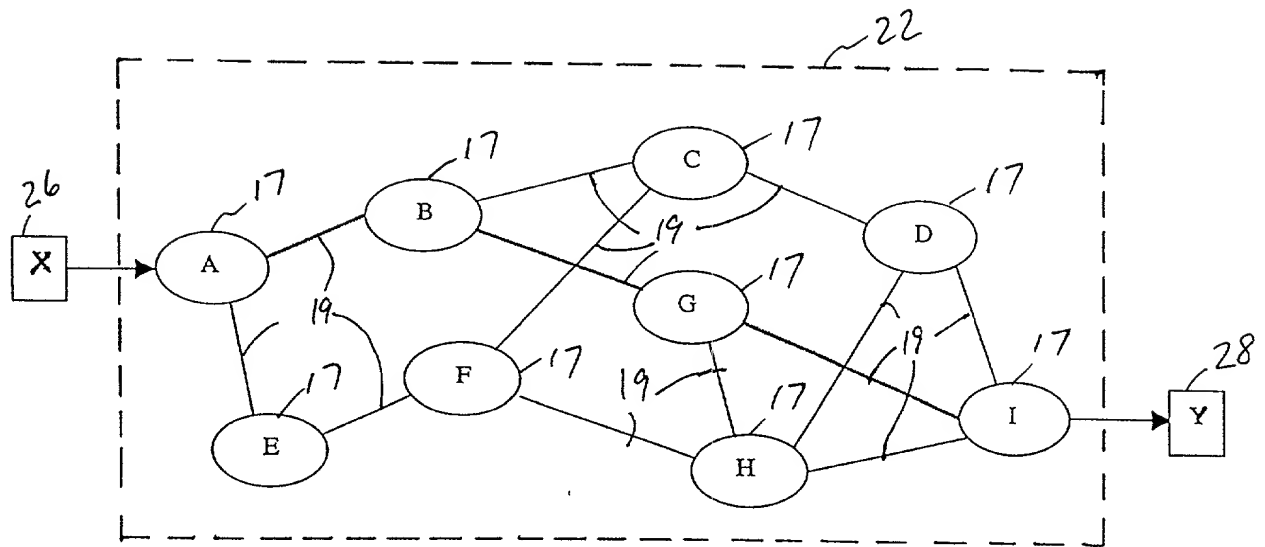
FIG. 2
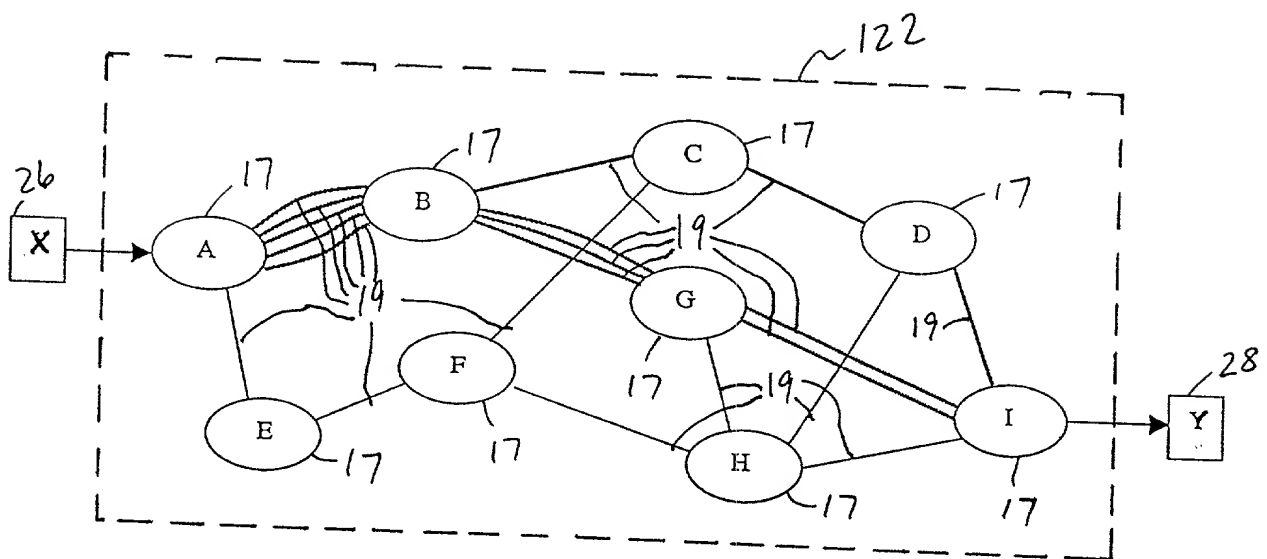
FIG.3

	DECLARATION, PETITION AND POWER OF ATTORNEY FOR PATENT APPLICATION	Attorney Docket No:
		IBNR-014

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

APPARATUS AND METHOD FOR FORWARDING DATA ON MULTIPLE LABEL-SWITCHED DATA PATHS

the specification of which (check only one):

___X___ is attached hereto.

_____ was filed as United States Patent Application

Serial No. _____

on _____

and was amended

on _____.

(if applicable)

_____ was filed as PCT Patent Application

Serial No. _____

on _____

and was amended under PCT Article 19

on _____.

(if applicable)
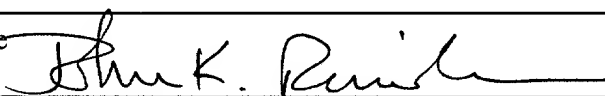
I hereby state that I have reviewed and understand the contents of the specification, including the claims as amended by any amendment referred to herein.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate or of any PCT international application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application(s) for patent or inventor's certificate or any PCT international application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) of which priority is claimed:

PRIOR FOREIGN/PCT APPLICATION(S) AND ANY PRIORITY CLAIMS UNDER 35 U.S.C. § 119:

COUNTRY (if PCT indicate PCT)	APPLICATION NUMBER	DATE OF FILING (day, month, year)	PRIORITY CLAIMED UNDER 35 U.S.C. § 119 (YES/NO)

DECLARATION, PETITION AND POWER OF ATTORNEY FOR PATENT APPLICATION	Attorney Docket No: **IBNR-014**

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States provisional application(s) listed below.

PRIOR U.S. APPLICATIONS FOR BENEFIT UNDER 35 U.S.C. § 119(e):

APPLICATION NUMBER	FILING DATE

I hereby claim the benefit under Title 35, United States Code, § 120 of any United States application(s) or PCT international application(s) designating the United States of America that is/are listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in that/those prior application(s) in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, § 1.56 which occurred between the filing date of the prior applications and the national or PCT international filing date of this application:

PRIOR U.S. APPLICATIONS OR PCT INTERNATIONAL APPLICATION(S) DESIGNATING THE U.S. FOR BENEFIT UNDER 35 U.S.C. § 120:

APPLICATION NUMBER (if PCT indicate PCT)	DATE OF FILING (day, month, year)	STATUS: (PATENTED, PENDING OR ABANDONED)

POWER OF ATTORNEY: As a named inventor, I hereby appoint the following attorneys and/or agents to prosecute this application and transact all business in the Patent and Trademark Office connected therewith.

Mark G. Lappin	Reg. No. 26,618	Steven M. Mills	Reg. No. 36,610
Toby H. Kusmer	Reg. No. 26,418	Anthony P. Onello, Jr.	Reg. No. 38,572
Robert J. Schiller	Reg. No. 19,248	Ronald R. Demsher	Reg. No. 42,478
Elizabeth A. Levy	Reg. No. 34,375	Carolyn G. d'Agincourt	Reg. No. 43,327
David Silverstein	Reg. No. 26,336	David M. Mello	Reg. No. 43,799

Send Correspondence to:	Direct Telephone Calls to:
Steven M. Mills, Esq. LAPPIN & KUSMER, LLP 200 State Street Boston, Massachusetts 02109	Steven M. Mills, Esq. (617) 330-1300 (617) 330-1311 (facsimile)

Wherefore I petition that letters patent be granted to me for the invention or discovery described and claimed in the attached specification and claims, and hereby subscribe my name to said specification and claims and to the foregoing declaration, power of attorney, and this petition.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Signature *John K. Renwick*	Date 9/16/1999

Full Name of 1st Inventor	Family Name **Renwick**	First Given Name **John**	Second Given Name **K.**
Residence & Citizenship	City **Minneapolis**	State or Foreign Country **Minnesota**	Country of Citizenship **U.S.A.**
Post Office Address	Post Office Address **3037 Holmes Avenue, Apt. 2**	City **Minneapolis**	State & Zip Code/Country **MN, 55408, U.S.A.**

DECLARATION, PETITION AND POWER OF ATTORNEY FOR PATENT APPLICATION		Attorney Docket No: **IBNR-014**
Signature _Ross W Callon_	Date _9/20/99_	

Full Name of 1st Inventor	Family Name **Callon**	First Given Name **Ross**	Second Given Name **W.**
Residence & Citizenship	City **Westford**	State or Foreign Country **Massachusetts**	Country of Citizenship **U.S.A.**
Post Office Address	Post Office Address **11 Applewood Drive**	City **Westford**	State & Zip Code/Country **MA, 01886, U.S.A.**